第 15 講　データの分析

PART1　データの整理と代表値

要点 ①　データの整理

「変量」… 身長，気温，テストの点数のように，人や物の特性を数量的に表すもの

「データ」… 変量の観測値や測定値などの集まり

データを整理する方法として，「度数分布表」や「ヒストグラム」がある。

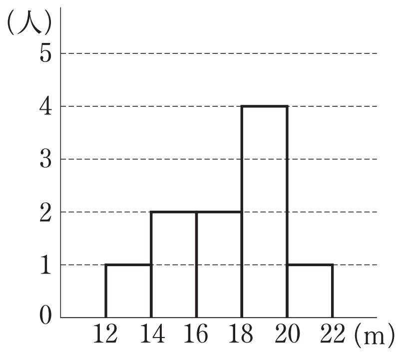
例　次のデータは，あるクラスの生徒 10 人のハンドボール投げ (m) の結果である。

13，21，15，18，18，16，17，15，19，18 (m)

度数分布表

階級 (m)	度数 (人)
12 以上 14 未満	
14 ～ 16	
16 ～ 18	
18 ～ 20	
20 ～ 22	
計	10

ヒストグラム



階級の幅は　　m

14 m 以上 16 m 未満の階級の相対度数は　　= 0.2

要点 ②　データの代表値

「代表値」… データ全体の特徴を 1 つの数値で表したもの

主な代表値として「　　　　」「　　　　」「　　　　」がある。

「平均値」… 変量 x についてのデータが，n 個の値 x_1，x_2，……，x_n であるとき，

これらの和を値の個数 n で割ったもの。「　」で表す。

$$\overline{x}=\frac{1}{n}(x_1+x_2+\cdots\cdots+x_n)$$

「中央値」… データを値の大きさの順(小さい順，大きい順)に並べたとき，
中央の位置にくる値。メジアンともいう。
データの個数が偶数のときは，中央に並ぶ 2 つの値の平均値。

「最頻値」… データにおいて，最も個数の多い値。モードともいう。
データが度数分布表に整理されているときは，度数が最も大きい
階級の階級値(階級の中央の値)を最頻値とする。

例 データ

13，21，15，18，18，16，17，15，19，18 (m)

について，代表値を求める。

このデータの平均値は

$$\overline{x}= \qquad\qquad = \frac{170}{10} = 17.0 \text{ (m)}$$

このデータを小さい順に並べると

13，15，15，16，17，18，18，18，19，21

よって，このデータの中央値は　　　　＝17.5 (m)

このデータの最頻値は　　m

※右表のように，度数分布表に整理されているとき，

最頻値は　　　　＝19 (m)

＜度数分布表＞

階級 (m)	度数 (人)
12 以上 14 未満	1
14 ～ 16	2
16 ～ 18	2
18 ～ 20	4
20 ～ 22	1
計	10

平均値は極端に離れた値の影響を受けやすく，中央値は極端に離れた値の影響を受けにくい。

1 (1) 次のデータの平均値，中央値をそれぞれ求めよ。

(i) 32，40，31，45，80

(ii) 45，41，39，50，44，45

(2) 右の表は，あるクラスの生徒27人のソフトボール投げの結果である。

(i) 中央値が含まれる階級を求めよ。

(ii) 最頻値を求めよ。

階級 (m)	度数 (人)
12 以上 14 未満	2
14 ～ 16	4
16 ～ 18	9
18 ～ 20	5
20 ～ 22	7
計	27

Point Pickup

代表値（平均値，最頻値，中央値）

平均値 … $\overline{x}=\dfrac{1}{n}(x_1+x_2+\cdots\cdots+x_n)$

平均値は，極端に離れた値の影響を受けやすい

最頻値 … 最も個数の多い値。

度数分布表に整理されているときは，度数が最も大きい階級の階級値。

中央値 … 大きさの順に並べたとき，中央の位置にくる値。

データの個数が偶数のときは，中央に並ぶ2つの値の平均値。

中央値は，極端に離れた値の影響を受けにくい。

第15講　PART1 確認問題

1　次の８個のデータがある。

1，2，4，5，5，8，10，13

平均値は［ア］である。

2　次のような(A)，(B) 2 つのデータがある。

(A)　1，2，5，8，9，11，13

(B)　3，3，6，7，11，14，14，15

(A)の中央値は［イ］，

(B)の中央値は［ウ］である。

3　右のような度数分布表で表される 25 個のデータがある。

中央値が含まれる階級は

［エ］以上［オ］未満

である。

階級	度数
以上　未満	
1 ～ 3	2
3 ～ 5	4
5 ～ 7	7
7 ～ 9	9
9 ～ 11	3
計	25

4　右のような度数分布表で表される 25 個のデータがある。

最頻値は［カ］である。

階級	度数
以上　未満	
1 ～ 3	2
3 ～ 5	4
5 ～ 7	7
7 ～ 9	9
9 ～ 11	3
計	25

第 15 講　データの分析

PART1　データの整理と代表値

1　(1)　次のデータの平均値，中央値をそれぞれ求めよ。

(i)　32，40，31，45，80

(ii)　45，41，39，50，44，45

(2)　右の表は，あるクラスの生徒 27 人のソフトボール投げの結果である。

(i)　中央値が含まれる階級を求めよ。

(ii)　最頻値を求めよ。

階級 (m)	度数 (人)
12 以上 14 未満	2
14 ～ 16	4
16 ～ 18	9
18 ～ 20	5
20 ～ 22	7
計	27

[解答]

(1)(i)　平均値は $\dfrac{32+40+31+45+80}{5}=\dfrac{228}{5}=45.6$

データを小さい順に並べると 31，32，40，45，80

よって，中央値は 40

(ii)　平均値は $\dfrac{45+41+39+50+44+45}{6}=\dfrac{264}{6}=44$

データを小さい順に並べると 39，41，44，45，45，50

よって，中央値は $\dfrac{44+45}{2}=44.5$

(2)(i)　度数の合計が 27 人より，中央値は値の小さい方から(大きい方から)14 番目の値である。これが含まれる階級は 16m 以上 18m 未満

(ii)　度数が最も大きい階級は 16m 以上 18m 未満。よって，最頻値は $\dfrac{16+18}{2}=17$ (m)

第15講　PART1 確認問題 解答

1

[正解] ア：6

[解説]

平均値は，[A]

$$\frac{1+2+4+5+5+8+10+13}{8}$$

$$=\frac{48}{8}=6$$

> [A]　平均値
> $\overline{x}=\frac{1}{n}(x_1+x_2+\cdots+x_n)$

2

[正解] イ：8，ウ：9

[解説]

(A)　データが7個だから，値の小さい方から4番目にくる値が中央値になる。[B]

中央値は，8

(B)　データが8個だから，小さい方から4番目と5番目にくる値の平均が中央値となる。[C]

中央値は，$\frac{7+11}{2}=9$

> [B]　中央値…大きさの順に並べたとき中央にくる値

> [C]　データの個数が偶数のときは，中央の2つの値の平均値が中央値となる。

3

[正解] エ：5，オ：7

[解説]

中央値は,値の小さい方から13番目にくる値である。

[D]

これが含まれる階級は**5以上7未満**。

> [D]　25個の中央にくる値は，小さい方から13番目である。

4

[正解] カ：8

[解説]

個数が最も多い階級は7以上9未満である。

最頻値は，この階級の階級値だから，[E]

$$\frac{7+9}{2}=8$$

> [E]　度数分布表から最頻値を求めるときは，度数が最も大きい階級の階級値を求める。